

Sensitivity of projected climate impacts to climate model weighting: multi-sector analysis in eastern Africa

Article (Published Version)

Kolusu, Seshagiri Rao, Siderius, Christian, Todd, Martin C, Bhawe, Ajay, Conway, Declan and James, Rachel (2021) Sensitivity of projected climate impacts to climate model weighting: multi-sector analysis in eastern Africa. *Climatic Change*, 164 (a36). ISSN 0165-0009

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/97869/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.



Sensitivity of projected climate impacts to climate model weighting: multi-sector analysis in eastern Africa

Seshagiri Rao Kolusu^{1,2} • Christian Siderius^{3,4} • Martin C. Todd² • Ajay Bhawe⁵ • Declan Conway³ • Rachel James⁶ • Richard Washington⁶ • Robel Geressu⁷ • Julien J. Harou^{7,8} • Japhet J. Kashaigili⁹

Received: 3 July 2020 / Accepted: 6 January 2021 / Published online: 10 February 2021
© The Author(s) 2021

Abstract

Uncertainty in long-term projections of future climate can be substantial and presents a major challenge to climate change adaptation planning. This is especially so for projections of future precipitation in most tropical regions, at the spatial scale of many adaptation decisions in water-related sectors. Attempts have been made to constrain the uncertainty in climate projections, based on the recognised premise that not all of the climate models openly available perform equally well. However, there is no agreed ‘good practice’ on how to weight climate models. Nor is it clear to what extent model weighting can constrain uncertainty in decision-relevant climate quantities. We address this challenge, for climate projection information relevant to ‘high stakes’ investment decisions across the ‘water-energy-food’ sectors, using two case-study river basins in Tanzania and Malawi. We compare future climate risk profiles of simple decision-relevant indicators for water-related sectors, derived using hydrological and water resources models, which

✉ Seshagiri Rao Kolusu
seshagirirao.kolusu@metoffice.gov.uk

¹ Met Office, FitzRoy Road, Exeter EX1 3PB, UK

² Department of Geography, University of Sussex, Brighton BN1 9QS, UK

³ Grantham Research Institute on Climate Change and the Environment, London School of Economics, Houghton Street, London WC2A 2AE, UK

⁴ Uncharted Waters Research, Sydney, Australia

⁵ School of Engineering, Newcastle University, Newcastle NE1 7RU, UK

⁶ Climate Research Laboratory, Oxford University Centre for the Environment, Oxford, UK

⁷ Department of Mechanical, Aerospace and Civil Engineering, University of Manchester, Manchester, UK

⁸ Department of Civil, Environmental and Geomatic Engineering, University College London, London, UK

⁹ Sokoine University of Agriculture, Morogoro, Tanzania

are driven by an ensemble of future climate model projections. In generating these ensembles, we implement a range of climate model weighting approaches, based on context-relevant climate model performance metrics and assessment. Our case-specific results show the various model weighting approaches have limited systematic effect on the spread of risk profiles. Sensitivity to climate model weighting is lower than overall uncertainty and is considerably less than the uncertainty resulting from bias correction methodologies. However, some of the more subtle effects on sectoral risk profiles from the more ‘aggressive’ model weighting approaches could be important to investment decisions depending on the decision context. For application, model weighting is justified in principle, but a credible approach should be very carefully designed and rooted in robust understanding of relevant physical processes to formulate appropriate metrics.

Keywords Climate modelling, model weighting · Impact modelling and water-food-energy nexus

1 Introduction

Climate change adaptation efforts are often informed by projections from climate models, e.g. those from the Climate Model Intercomparison Project (CMIP) multi-model ‘ensemble’ conditioned by a set of common scenarios of anthropogenic greenhouse gas and other radiative forcings. There is a growing ‘market’ of portals providing such information (see examples in Nissan et al. 2019; Hewitson et al. 2017). It is well understood by climate scientists, but not always effectively captured or communicated to other researchers and decision-makers, that there is considerable uncertainty in projected future climate, especially at finer space/time scales, for longer time horizons, and for some climate variables (e.g. rainfall) more than others, particularly in the tropics (IPCC 2013; Knutti 2008). Projection uncertainty is a challenge for decision makers and hinders the incorporation of climate risks into long-term strategic development plans.

There are now increasing attempts by climate scientists to understand and to constrain projection uncertainty, as climate models in the available multi-model ensembles are not equally good, nor are they truly independent of each other (Knutti et al. 2010; Masson and Knutti 2011; Sanderson et al. 2017; Pennell and Reichler 2011; Massoud et al. 2019, 2020), despite an acceptance of ‘model democracy’ in IPCC reports. Approaches have therefore been proposed to discriminate between climate models in terms of their ‘trustworthiness’ (see Eyring et al. 2019 for a review and examples), based on how well the models represent key climate variables (Baumberger et al. 2017), i.e. their ‘fit’ to observations. The logic being that models which simulate today’s climate well are likely to have more plausible projections of future change, which assumes processes causing present-day biases and errors are a cause of future climate uncertainty. Some studies have been able to exclude projected futures (e.g. Munday and Washington 2019) or constrain uncertainty (Gershunov et al. 2019) on the basis of present-day process biases. However, other studies find that excluding the ‘poorest’ models has little effect on the range of uncertainty (e.g. Déqué and Somot 2010; Knutti et al. 2010; Rowell et al. 2016). Despite the growing acceptance of the desirability of distilled or refined climate ensemble information, there remains no accepted or agreed set of procedures to measure model performance and to weight accordingly.

At the same time, Decision Making Under Climate Uncertainty (DMUCU) methods are being developed that explicitly recognise and respond to the existence of deep uncertainty (see examples in Marchau et al. 2019) and are being applied (e.g. Ray and Brown 2015; Hurford et al. 2020), although this is at present rare in low-income countries (Bhave et al. 2018). Some

(e.g. Nissan et al. 2019; Weaver et al. 2013) argue that climate models are not fit for the purpose of directly informing many local adaptation decisions, and some DMUCU approaches avoid reliance on climate model output, instead aiming to identify decisions that are robust to wide ranges of possible futures. However, most DMUCU methods use climate models at some stage of the analysis (Hunt and Watkiss 2011). Weighted climate model projections might arguably be useful for DMUCU approaches (e.g. Erfani et al. 2018) although few applications currently exist.

This paper aims to inform this evolving discussion by establishing to what extent weighting approaches to constrain climate uncertainty might affect the nature of decision-relevant climate impact information. We provide two case studies (from the Future Climate For Africa (FCFA) project ‘UMFULA’, <https://futureclimateafrica.org/project/umfula/>); the Rufiji River in Tanzania and the Lake-Malawi-Shire River (Fig. 1), which are characteristic of the challenges around sustainable and climate-resilient development across the ‘water-energy-food’ (WEF) nexus (Section 2.3.2). This is especially pertinent given the magnitude of current and proposed major (‘high stakes’) investments in agriculture and hydropower (Bhave et al. 2020; Conway et al. 2019; Luhunga et al. 2018). For these case studies, we did the following:

- (1) Assess a range of approaches to determine climate model ‘trustworthiness’ which we translate into model ‘weightings’ (Section 2.1), used in climate impact assessment.
- (2) Assess the effect of these climate model weighting approaches on projected risk analysis and decision-relevant water resource indicators and the associated uncertainty (Section 2.2).

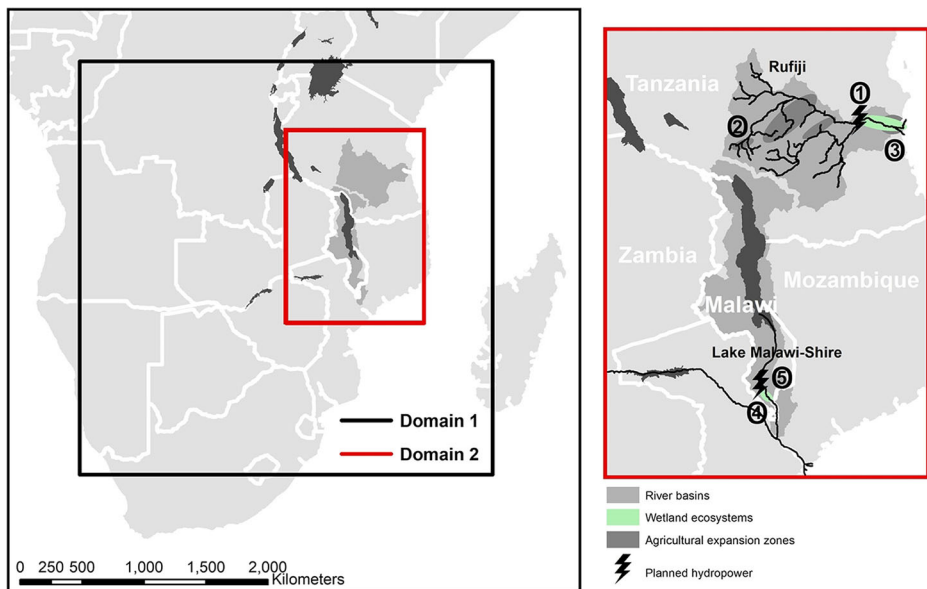


Fig. 1 Location of study sites. Left panel; the black and red boxes represent the primary sub-continental and secondary regional domains, respectively, over which climate model performance is evaluated. The right panel shows the locations of water resource developments in the two case study basins, with (1) Julius Nyerere Hydropower Project (JNHP), (2) upstream rice areas in the Southern Agricultural Growth Corridor of Tanzania (SAGCOT), (3) Rufiji delta ecosystem, (4) Elephant marshes, and (5) Kholombidzo Hydropower and Shire Valley Transformation Project (SVTP)

We then assess to what extent water resource management decisions might be influenced by the choice of climate model weighting approach before considering the implications for the broader issue of whether and/or how to adopt model weighting in adaptation.

2 Methods

The outcome of our analysis is an assessment of specific climate risks to sectoral water resources, and its sensitivity to climate model weighting, using the method below (Fig. 2).

2.1 Climate model scenarios and observations (step 1 in Fig. 2)

We use 32 of the CMIP5 models (Taylor et al. 2012) (Table S1), with all the data necessary to support the calculation of the surface water balance to drive the hydrological impact models (Section 2.3). We use simulations of the historical period and over the period 2020–2050 derived using the Representative Concentration Pathway RCP8.5 forcing scenario. For models with more than one ensemble member, we use only the first member. Model weighting (Section 2.2) is derived via comparison with historical observations: The Global Precipitation Climatology Centre (GPCC) monthly precipitation version 7 (Schneider et al. 2017); and CRU Temperature V4.02 (Harris et al. 2013). All model and observed data were remapped to a common grid of $1.0 \times 1.0^\circ$, roughly equivalent to that of the highest resolution CMIP5 models.

2.2 Climate model weighting (step 2 in Fig. 2)

We use four model weighting methods (W1–W4 below), broadly representative and indicative of those developed in the literature. The model weightings are applied in step 4 of our analysis to determine climate impact risk profiles (Section 3.2), and we also apply the weights to the distribution of projected changes in rainfall and in river flow (Section 3.1.2). We make no judgement here about the relative merits of the weighting approaches, nor recommendations about the ‘best’ method.

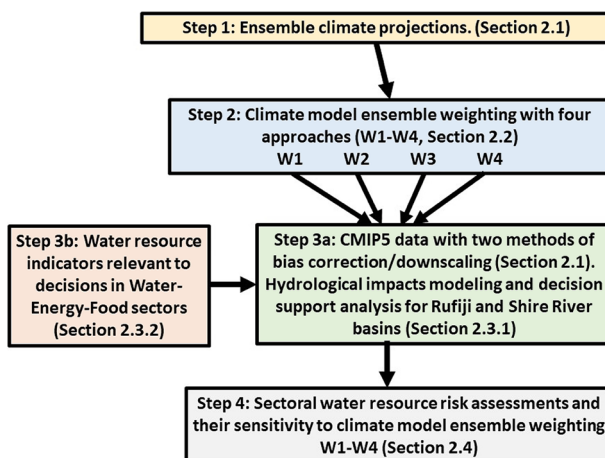


Fig. 2 Schematic of the major steps in the analysis

W1. Uniform weighting Each model is assumed to be equally trustworthy, effectively the default approach used in most impact assessment analyses.

W2. Binary inclusion/exclusion weighting by performance Models are assigned a weighting of either one or zero depending on their performance rank. Models are ranked from (1–32) according to a number of performance metrics (see below and Table 1). The average rank across multiple metrics is then derived. Once ranked, the top 50% of the 32 models are selected and assigned a weight of one. All other models deemed ‘unacceptable’ are weighted zero. This is akin, for example, to the ensemble sub-setting approach developed by Rowell et al. (2016) or by Gershunov et al. (2019).

W3. Model weighting by performance and independence Using the approach of Sanderson et al. (2017) utilised in the US fourth National Climate Assessment, each model is assigned a weight, incorporating a ‘skill’ weight, i.e. the model performance with respect to observations across the performance metrics in Table 1, and a model independence weight, i.e. the model performance with respect to all other models. The Sanderson et al. (2017) method is summarised in Section S1.

We assess models’ ‘fit for purpose’ by selecting performance metrics that are logically ‘problem-relevant’ (Baumberger et al. 2017; Eyring et al. 2019; Massoud et al. 2020), for the assessment of climate impacts on large-scale surface water balance for hydrological modelling. Accordingly, for the variables temperature (T) and precipitation (P), we define indicators of (i) mean state, including seasonality; (ii) variability, including both the magnitude and the dominant driving process, in this case, simplified to the teleconnection with the El Niño Southern Oscillation (ENSO) that exerts a dominant influence on the region (e.g. Kolusu et al. 2019; Harrison et al. 2019); and (iii) change; specifically recent multi-decadal trends when anthropogenic climate change is most pronounced. This is consistent with the recommendations for adaptation-focussed model evaluation of Nissan et al. (2019). However, we recognise that model performance over the larger scale and for a range of more fundamental metrics, like radiation budget, or global/regional sea surface temperature biases could be equally relevant to regional climate.

Each indicator is derived at the grid cell scale over a study domain centred on our region and large enough to capture the large-scale climate structure (15–45 °E, 0–30 °S; see the black box in Fig. 1). Then, we derive the performance metric as the spatial root mean squared error (RMSE) of the model versus observed values over the domain. Note that the RMSE of each model versus every other model is also used in establishing the independence weight in weighting method W3 (see Section S1).

Table 1 Climate indicators used in model weighting methods W2 and W3. These are derived for temperature and precipitation, for OND and JFM seasons. The performance metric is then defined as the RMSE of each indicator versus observations across the domain (and with respect to all other models for the independence weight in method W3, see Section S1)

Climate characteristic	Indicator
Mean state	Climatological mean (1976–2005) of T and P
Variability	Coefficient of variance (1850–2005) of T and P
Drivers of variability	Correlation of seasonal T, P values with Niño3.4 SST index
Trend	Linear trend (1976–2005)

Note that for methods W2 and W3, we explore the sensitivity of model ranks and weights to assumptions made in their derivation, specifically the sensitivity to the size of the study domain, the choice of performance metrics and uncertainty in observational data used. The results of this analysis are presented in full in supplementary material Section S3 and considered where relevant in Section 3.1.

W4. Model outlier plausibility weighting In this approach, the ‘outlier’ models are identified, i.e. those models whose precipitation climate change projections are most extreme relative to the rest of the ensemble: i.e., they are at the edges of the range of projections and they have the largest wetting or drying responses. These outliers are then assessed by expert judgement as to their plausibility (see Section S1). A weighting of zero is applied to models deemed not plausible, and this represents a bespoke, project-specific approach to model weighting. The underpinning rationale is that adaptation decisions, based on either a multi-model mean of climate projections or including the full ensemble, are likely to be heavily influenced by outlier models. Adaptation plans that are robust to climate change uncertainty will tend to be more expensive if the sampled uncertainty is skewed by outliers. It is therefore reasonable to assess whether such outliers are plausible. The approach follows from the recommendations of Rowell and Chadwick (2018) and exemplified in, e.g. Rowell (2019). The expert assessment suggested there is a basis for excluding the following six models: IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC-ESM, GISS-E2-H, GISS-E2-R, GISS-E2-R-CC, which are dominated by models from two ‘families’ of the model, IPSL and GISS (see Section S1).

2.3 Hydrological and water resources impact modelling (step 3 in Fig. 2)

2.3.1 Step 3a: hydrological models and driving climate scenario data

We explore the range of climate impacts on water resources and associated decision-making using hydrological and water resource models, specifically a basin-scale implementation of the LPJml for the Rufiji River and the Water Evaluation And Planning (WEAP) model for the Lake Malawi-Shire River case study. Characteristics of the study sites, including the decision context and the hydrological modelling are described in Siderius et al. (2018) for the Rufiji River and in Bhave et al. (2020) for the Lake Malawi-Shire River.

In each case, calibrated hydrological models are forced with multiple future climate scenarios (for the period indicative of 2021–2050) from the CMIP5 models (described in Section 2.1). The climate forcing fields include precipitation and variables required to derive evapotranspiration. We use two distinct methods to bias correct these climate model variables, indicative of end members of the spectrum of complexity in approaches widely used in the literature (Teutschbein and Seibert 2012; Seaby et al. 2013).

- (i). Change ‘delta’ factors. A ‘delta’ factor is derived for each climate model for each calendar month, defined as the ratio of the model’s mean future (2021–2050) to mean historical (1976–2005) value. The observed historical climate data is then perturbed using the ‘delta’ factors derived for all variables required to drive the hydrological

models, except temperature (used to derive evapotranspiration), which is calculated as the absolute change to generate future climate projections.

- (ii). Full quantile mapping (QM) bias correction (analysis conducted under the FCFA project AMMA-2050, Famien et al. 2018) in which the probability distribution function of the projected daily values for each climate variable is corrected based on a quantile mapping transfer function (described in detail in the supplementary section S2). This transfer function is derived from comparison of the modelled and observed daily probability distribution over the historical period. As such, the future climate scenarios incorporate the model's projected changes in both mean climate and variability at all time scales up to the decadal.

These approaches are applied to all variables required for the calculation of the water balance, specifically: precipitation, maximum and minimum air temperature, specific humidity, wind speed and downward shortwave radiation.

2.3.2 Step 3b: water resource development scenarios

A single-water resource development scenario was implemented for each basin based on existing sectoral development plans from national governments, such that it represents ambitious but plausible development involving 'high stakes' investment decisions. (Note that (i) no climate adaptation is assumed in the development scenarios and (ii) analysis of water resources under a complex range of various development scenarios and water management portfolio options is explored in a companion paper, Geressu et al. 2020). The Rufiji basin (Fig. 1) is earmarked for extensive development of water resources, notably (i) expansion of irrigated agriculture, through the Southern Agricultural Growth Corridor of Tanzania initiative (SAGCOT www.sagcot.co.tz), which will potentially increase irrigation water demand threefold (WREM International 2015)—this is represented in the hydrological model by adjusting current land use to the upper limit of planned irrigated area expansion and related water demand for 2035, as reported in the Rufiji Water Resources Management plan (WREM 2015), and (ii) expansion of hydropower, particularly the recently approved Julius Nyerere Hydropower Project (JNHP), a dam located at Stiegler's Gorge, close to the basin's delta (Fig. 1). JNHP will be one of the largest hydropower dams in Africa, with a generating capacity of 2115 MW. In the hydrological model, we use a simplified schematization and parameterisation of the JNHP and its operating rules and assume its main aim to be a guaranteed firm energy supply. We assume static operation rules, with no capacity to anticipate future inflows, which although unrealistic, ensures a consistent baseline to assess climate risk independent of adaptation measures.

The Lake Malawi-Shire River Basin is part of the Zambezi basin and covers most of Malawi (Fig. 1). Lake Malawi outflows into the Shire River, whose flows sustain more than 90% of Malawi's electricity production (through hydropower), irrigation and environmental flows for a Ramsar designated wetland; the Elephant Marsh. Lake Malawi outflows are regulated by the Kamuzu Barrage (see Bhawe et al. 2020 for further details). Relevant planned future developments include (i) a fourth hydropower station on the Shire River—the 200 MW Kholombidzo hydroelectric power station, and (ii) the Shire Valley Transformation Project which seeks to more than double the present area under irrigation. We incorporate these future water demand plans into the WEAP water resources model, in terms of average monthly river flow. The Kamuzu Barrage operating rules are assumed to be static (based on current practice).

2.4 Assessing climate risk to water resource development (step 4 in Fig. 2)

To explore potential risks posed by climate change to water resource development, we define a set of indicative decision and planning-relevant water resource indicators. These are minimum performance requirements across the three main WEF sectors (Table 2) informed by consultation with stakeholders (Bhave et al. 2020; Geressu et al. 2020). Specifically the following:

- (ii) For the water sector, minimum environmental river flows to sustain ecological function;
- (iii) For the energy sector, river flows to sustain hydropower generation;
- (iv) For the agriculture/food sector, river flows to maintain irrigated agricultural production.

All these performance indicators are expressed in terms of the percentage of time, either in months or years, that a specific performance criterion is not met during the future simulation period (2020–2050). This, we call the ‘rate of failure’ for each indicator, and it is derived for (i) the observational period and (ii) the future (2020–2050), from the hydrological simulation forced with the ensemble of climate model scenarios. The latter gives a distribution of future rates of failure, which we refer to as the future ‘risk profile’, and we then apply each climate model’s weight (under W1–W4) to the risk profile (described in detail in Section 3.2).

The result for each basin and sector is four distinct climate risk profiles for the future epoch, each being a climate model-weighted distribution of rates of failure for the sectoral performance metric (analysed in Section 3.2). Note that we caution against the interpretation of these as probability distributions in the formal sense as they do not represent the true likelihood of occurrence, especially given the limited nature of the CMIP ‘ensemble of opportunity’, see e.g. Stainforth et al. (2007). Note also that the indicative sectoral performance indicators identified (Table 2) are simple translations of water quantity into sectoral ‘failure’ risk with no complex or compound effects included. No account is taken, for example, of the rules by which water is used in irrigation or reservoir operations; the potential flood risk to agriculture; combined risk, investment or operation optimisation and trade-offs; and the effects of adaptation policies and interventions in any of the sectors.

3 Results

3.1 Climate model weights and resulting climate projection uncertainty

3.1.1 Consistency in climate model weights

The spread of climate model ranks (from weighting method W2) and weights (from weighting method W3) across the ensemble are shown in Fig. 3a and b, respectively, and their association in Fig. 3c. Also shown for each model is the projected annual mean change in precipitation (dP) and basin runoff (dQ) for the Rufiji (Fig. 3) and Shire rivers (Fig. 3b) (the climate model weights and ranks are derived over a larger domain and are the same in both cases). The results indicate the following: First, the ordering of models by combined weight is very close to that by rank as expected, given the contribution of skill to the weighting (Figs. 3a–c), although some models do deviate, e.g. MICRO5 and FGOALS-g2 (Fig. 3a and b). Second, there is considerable consistency in, and relatively little averaging-

Table 2 Decision-relevant water resource risk indicators

Sector	Basin	Indicator	Description
Energy	Rufiji	RP1. Percentage of years with at least 1 month of no turbine outflow (NTF), because live storage is depleted, in a 30-year time period.	Hydropower no 'turbine flow' (NTF) occurs when NHP reservoir lake storage falls to a level which cannot guarantee the required monthly mean turbine flow of 687 m ³ /s to generate firm energy. This can be expressed as a risk of NTF occurring at least once a year. It reflects an irregular but still recurrent situation in which Tanzania requires access to backup generation facilities, which might be expensive given the expected size of the NHP and its contribution to overall electricity generation (Siderius et al. 2018).
	Shire	SP1. Months of inadequate Shire river flow. Percentage of months with Shire flows < 265 m ³ /s at the 3 hydropower stations on the Shire over a 30-year period.	Three run-of-the-river hydropower stations in the Shire River require approximately the same river flow of ~265 m ³ /s for peak hydropower production, which falls linearly with reducing discharge. Also, if Lake Malawi levels drop below the outflow threshold, reduced Shire river flow during non-rainy season would affect hydropower production. These are expressed as a simple percentage based on the number of months in a 30-year time period with flow less than ~265 m ³ /s (Bhave et al. in review).
Water (environmental flows)	Rufiji	RP2. Risk of not achieving at least 2500 m ³ /s flow for 1 month every second year.	Environmental flows downstream of the dam depend on sufficient peak discharge to refill the delta lakes and surrounding wetlands and to flood farmland to replenish nutrients and soil moisture, flush the delta and buffer salinisation, creating vital conditions for fisheries. Full-bank flow is reached at a threshold of 2500 m ³ /s. Historically, this is exceeded once every 2 years, and 3000 m ³ /s is exceeded every 4 years. Based on a (limited) analysis of daily flows in the early 1970s the duration of flows above this threshold was determined at, on average, 32 days, so one full month (Duvail and Hamerlynck 2007).
	Shire	SP2. Risk of no outflows from Lake Malawi changing seasonal flow dynamics in the Elephant Marsh	The Elephant Marsh wetland is located downstream of hydropower stations and is a designated Ramsar wetland of global biodiversity importance. However, no Environmental Flow Requirement (EFR) assessment for the wetland is available. Based on Shire River flow time series at Chiromo gauging station prior to the construction of the Kamuzu Barrage (1965), ~230 m ³ /s is the lowest monthly discharge in the 1960–1964 period. We therefore use 230 m ³ /s as a metric for EFR at the Elephant Marsh and calculate a simple percentage based on the number of months in a 30-year time period with flow less than ~230 m ³ /s (Bhave et al. in review).
Food	Rufiji	RP3. Number of months irrigation demand for current and proposed irrigation projects is not met.	Agriculture relies heavily on irrigation in several parts of the basin. Here, we plot the reliability of irrigation supply to the irrigated areas upstream of the Usangu wetlands where the largest rice farms are situated. There is hardly any infrastructure upstream to

Table 2 (continued)

Sector	Basin	Indicator	Description
Shire	SP3.	Number of months irrigation demand for current and proposed irrigation projects not met.	<p>store and release water, so agriculture is dependent on timely and regular rainfall (WREM 2015; Lankford and Beale 2007).</p> <p>The Shire Valley Transformation Project (SVTP) has been formulated with defined monthly irrigation water requirements, ranging from 37 to 50 m³/s to be diverted from the Shire River. This monthly requirement is used as a metric for irrigation (Bhave et al. in review).</p>

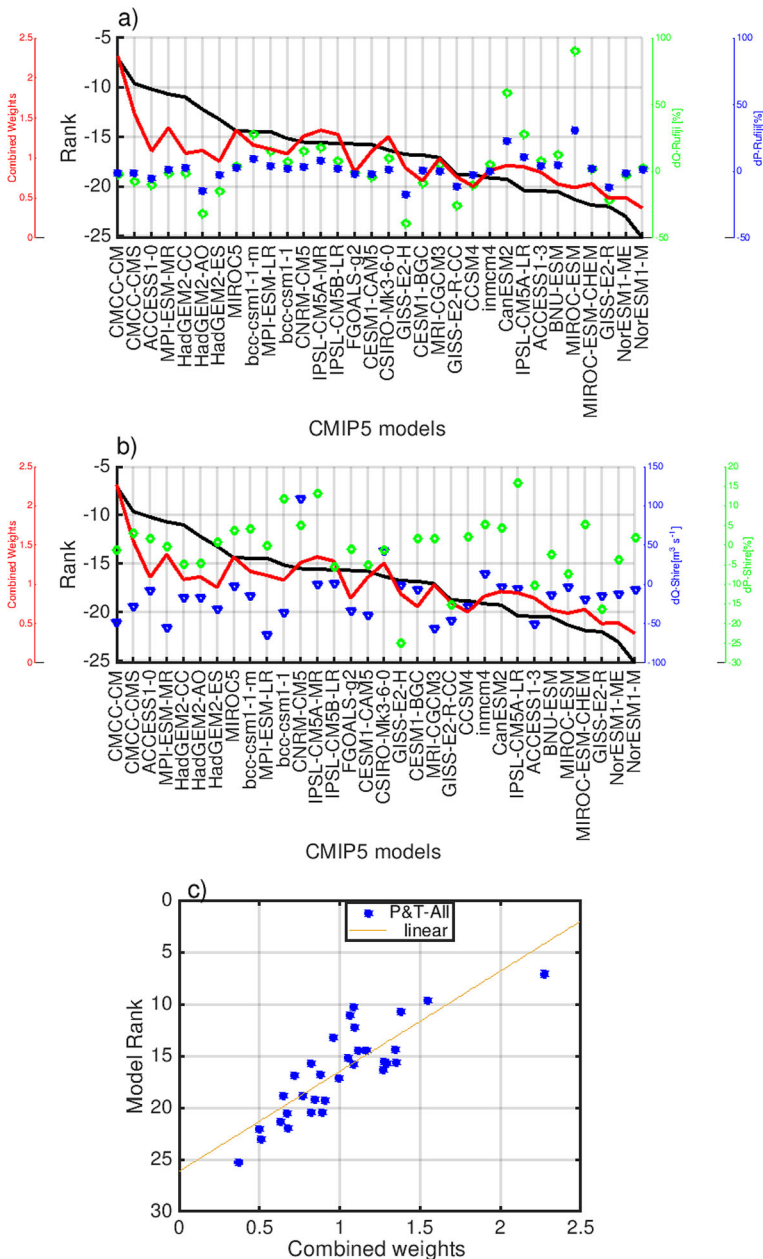


Fig. 3 **a–c** Climate model ranks and weights using methods W2 and W3. **a, b** Model ranks (ranks are multiplied by -1 for plotting purpose) from W2 (black line) and weights from W3 (red line) (both ordered left to right by W2 rank), with mean projected change (between 1975 and 2005 and 2021–2050) in precipitation (green diamonds) and simulated river flow (dQ , blue triangles) for **a** Rujiji and **b** Shire basins. **c** Relationship of climate model rank for method W2 (y -axis) and weights from W3 (x -axis)

out, in ranking across the 16 performance metrics with the highest and lowest ranked model (CMCC-CM and NORESM1-M, respectively) have mean ranks (across all performance

metrics) of 7 and 25 (out of 32). Third, W3 weights range from around 0.5 to 2 providing considerable potential to influence projected decision-relevant impacts. The magnitude of W3 weightings are quite sensitive to the number of metrics used (Section 3S3, Fig. S1b), which complicates the practical application of such approaches and we suggest that to obtain stable weights (W3), multiple performance metrics are used. Fourth, there is little indication of a simple and systematic relationship between the magnitude of projected future change and model performance. However, for the Rufiji basin, many of the models which have the strongest projected changes in both mean annual precipitation and resulting Rufiji River flow tend to be those with relatively low-performance ranking and weights. These include MIROC-ESM, CanESM, IPSL-CM5A-LR, and GISS-E2-H in Fig. 3, many of which are those excluded in the W4 method. Although this suggests a potential sensitivity of water resource risk profiles to model weighting, which we explore in Section 3.2, water resource risk is likely to be associated with reduced river flow, rather than the models showing strongly increased flow, which is most pronounced in Fig. 3.

The magnitude of model weights is quite sensitive to the choice of performance metrics (Fig. S1), and as such, the weighting sensitivity parameter in Eqs. (2) and (3) in supplementary material will require careful tuning to ensure an acceptable range of weights. Model ranks and weights are strongly correlated across the two small and large study domains (not shown). There is low sensitivity to uncertainty in observational data (Table S2).

3.1.2 Projected climate changes and mean hydrological impacts and associated uncertainty

We provide a first-order assessment of the effect of model weighting on projected hydrology by applying the weights to the ensemble of future changes in mean annual basin-averaged rainfall (dP) and simulated runoff (dQ), using the delta factor bias correction (Section 2.3.1(i)). (Note that these ensemble distributions of weighted dP and dQ are distinct from the weighted risk profiles analysed in Section 3.2). Figure 4 panels a and c show the CMIP5 model projections of precipitation for the Rufiji and Shire case-study regions for the October–March wet season. In both cases, the uniform weighting (W1) shows small CMIP5 multi-model mean projected changes for both basins with a marginal increase in mean annual precipitation of a few percent. Uncertainty is high however with the interquartile range (IQR) of the CMIP5 model ensemble including extending above and below the zero change line. Outlier models can show very substantial changes up to +32% and –17% for the Rufiji and +16% to –24% for the Shire. Model weighting does not change substantially the multi-model mean precipitation change projection in either case. However, the extreme ends of the inter-model distribution of projected change are constrained substantially, notably for W2 and W4 which involve model exclusion. As such, there is an indication that binary weighting by model performance can strongly moderate the range of plausible future climate, but without necessarily providing a clearer signal on the likely direction of future change. Assessment of the sensitivity in the projected uncertainty to weighting scheme options is provided in Section 3S3, Table S2. The choice of performance metrics shows considerable influence on uncertainty, and others are minor.

Projected changes in a 30-year mean annual river flow (dQ, Fig. 4 panels b and d) present very similar patterns of results. Note that in the Rufiji River basin, hydrology acts to magnify changes such that dQ is greater than dP; during the rainy season, soils remain close to saturation and most additional precipitation is transformed into additional runoff leading to a

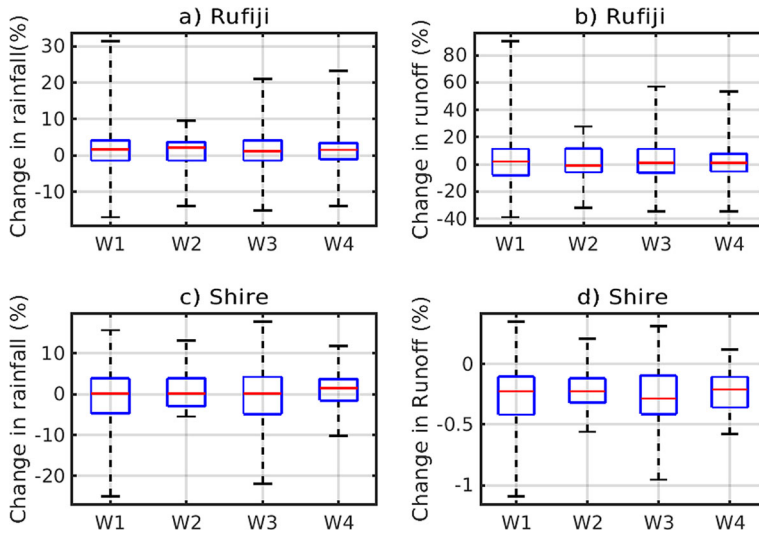


Fig. 4 Projected changes (2021–2050 minus 1976–2005) in basin-averaged mean annual rainfall (a, c) and river flow (b) and Lake Malawi level (d) for the two study sites after applying the four model weighting approaches to the CMIP5 ensemble. The box-whisker plots in each case show the distribution of projected changes across the CMIP5 multi-model ensemble. The red line shows the median value, blue box is 25th/75th percentiles and black whiskers are absolute model maximum and minimum. Values are the percentage change averaged over ONDJFM season. Performance weights/ranks are based on the domain 0–50 °E, 5–35 °S shown in Fig. 1)

magnified dQ response. As such, the effect of model weighting in modulating projected uncertainty is proportionally greater. The impact of these projected changes on decision-relevant metrics is considered in the next section.

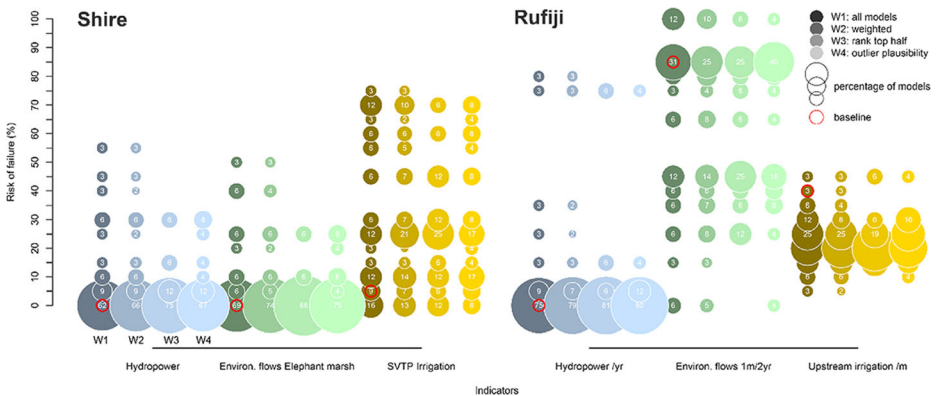


Fig. 5 Projected and present-day profiles of risk for hydropower, environmental flows and upstream irrigation demand indicators for Shire (left) and Rufiji (right). Y-axis is the sectoral indicator failure rate (percentage of months or years, see Table 2, in 5% risk percentage intervals). For each 5% interval of rate of failure, the model count in that interval is scaled by the weight for each model (then converted into a percentage of the total number of models). The red circle shows the present-day baseline risk. Circles show the failure rate emerging from the ensemble of CMIP climate models, in which the size of the circle is proportional to the percentage of models showing that the risk magnitude and the precise percentage of models indicated in the centre of each circle. Results are derived using the method shown in Fig. 2 and described in Section 2. CMIP5 projections use the delta factor climate model bias correction approach (Step 3a, Section 2.2.1)

3.2 Projected changes in decision-relevant metrics across the WEF sectors

Figures 5 and 6 show the weighted risk profiles, i.e. the spread in sectoral risk associated with different climate change projections (with the risk in the present-day baseline period also indicated). These risk profiles are derived by calculating, for each climate model scenario, the rate of failure for each sectoral indicator for water (environmental flows), energy (hydropower generation) and food (agricultural production) (see Table 2 and Section 2.2.2), i.e. the percentage of months/years in which the minimum sectoral performance threshold (Table 2) was not met. The risk profiles show the number count (represented by the size of the circle) of model scenarios showing failure rates in 5% categories (0–5%, 5–10%, 10–15%...95–100%), with the count being weighted, i.e. a model with a weight of 2.0 contributes four times that of a model weighted 0.5 to the size of the circle in the risk profile. In this way, and because we are interested in risks to WEF decisions, the climate model weights (W1–W4) are applied to the risk profile of sectoral ‘failure’ across the model ensemble, emerging from the unweighted projected changes in river flows (and is thus distinct from the weighted changes to river flow values shown in Fig. 4b and d). Figures 5 and 6 differ in being derived using the simple ‘delta factor’ (DF) and the more sophisticated quantile mapping (QM) methods of climate model bias correction to drive the hydrological models (Step 3a, Section 2.2.1), respectively, and below, we highlight the sensitivity of projected risk to that step in the analysis chain.

Consider first, the baseline risks (red circles in Figs. 5 and 6). For the Rufiji River, we note a high baseline risk to environmental flows assuming the JNHP hydropower dam is completed, and upstream irrigation has expanded, as the size of the reservoir and need to preserve water would reduce peak flows for delta wetland flooding. For the Shire, the simulated risk to environmental flows for Elephant Marsh is low (~5%). The risk to irrigation supply is also high in both basins, with water shortages estimated to occur in ~45% months. The risk to hydropower generation is low in both basins (<5%). In the Rufiji, reservoir volume of the JNHP is sufficient to buffer variability in runoff as experienced over the last 30 years (but note that risks are much higher considering a longer historic period including droughts in the 1920s and 1930s, not shown, see Siderius et al. in review).

The projected climate change impacts on sectoral risks include a wide range of outcomes. We see the following:

Rufiji River Figure 5 (for the DF bias correction approach Step 3a Section 2.2.1) shows that risks to environmental flows downstream of the JNHP are broadly reduced with a long tail of reduced risk extending to zero. The mode of the projected risk profile is similar to the baseline. Only a small percentage of models show an increased risk. The effects of model weighting reduce the occurrence of projected increased risk. The risk profiles from the QM bias correction (Fig. 6) show higher uncertainty than for the DF analysis with greater plausibility of reduced risk and greater plausibility of increased risk. The effect of weighting is small and if anything, increases rather than reduces risk.

For upstream, irrigated agriculture projected climate change poses an overall reduced risk on average by roughly half to around ~25% and only a marginal indication of increased risk. This is because most models project wetter conditions over the upland western Rufiji basin where runoff supplying the irrigation sites is generated. The effects of model weighting on increased risk are small but the more ‘aggressive’ weighting methods W3 and W4 suggest that greatly reduced risk is not plausible, as a result of the exclusion of strongly wetting models. The results using the QM bias correction are similar (Fig. 6).

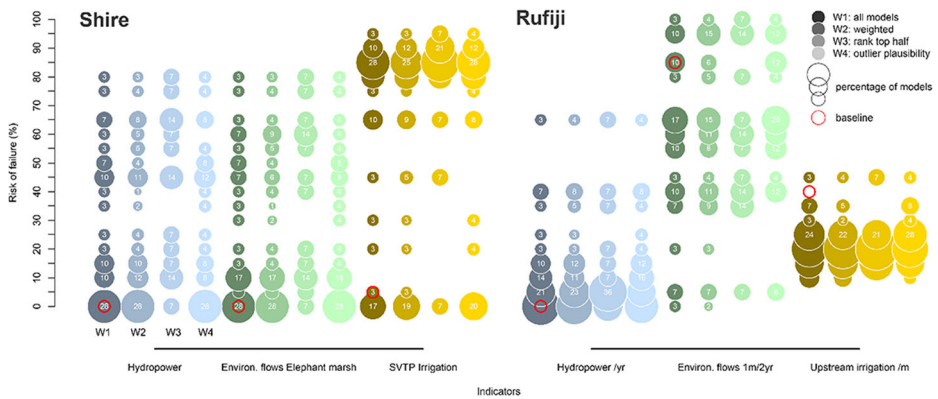


Fig. 6 The same as Fig. 5 but derived using the Quantile mapping climate model bias correction method (see Step 3a, Section 2.2.1)

The risk profiles for hydropower are highly skewed. Whilst the vast majority of models suggest a similar risk to today, there are a small number of models indicating dramatically increased failure rates. This is presumably as a result of projected basin-wide drying, indicating non-linearity between changes in mean rainfall and flow and decision-critical risks. This weak indication of high risk presents a challenge to investment decision making. A challenge that is not ameliorated by model performance weighting with the high magnitude risk remaining across all four weighted risk profiles. The results using the QM bias correction (Fig. 6) show greater plausibility of increased risk. Weighting has relatively little impact, although with some indication of a stronger increase in risk.

Shire River The already low risk to environmental flow is projected to reduce on average but with a small number of models indicating potentially increasing risk. The risk profiles for the ‘aggressive’ model weighting schemes with model inclusion tend to reduce this high-risk end of the profile consistent with Fig. 4c and d. The risk profiles for hydropower follow a similar pattern with a dominant reduction in the existing low failure risk, but with a small indication of dramatically increased risk, noticeably moderated by ‘aggressive’ model weighting.

For irrigated agriculture, the future risk profile shows increased risk on average with very considerable uncertainty reflecting the spread in model mean rainfall (Fig. 4). Model weighting appears to have a limited impact. Results with the QM bias correction (Fig. 6) show markedly different risk profiles. For environmental flows and hydropower, the risk profile is far less skewed with much greater plausibility of high risks. For irrigation, the distribution profile is very different and again, there is much higher plausibility of dramatically increased risk to the agriculture sector. Comparatively, climate model weighting has little impact.

Overall, there is relatively little sensitivity of the risk profiles to the various model weighting methods. Climate model performance weighting as implemented here has only a limited impact on projected risk uncertainty. This applies even in this case where the water resource indicators are sensitive to the extremes of the projected climate distribution, i.e. the risks across WEF sectors are most sensitive to future drying. Although several of the strongest wetting and drying climate projections are downweighted/excluded, none of the W2–W4 methods eliminates all of the more extreme projections; the risk of sectoral failure typically remains plausible. However, within these uncertain future risk profiles, subtle differences may provide more instructive information. For irrigation in the Rufiji, the two weighting sub-setting

methods (W2 and W4) show a reduced likelihood of lower risks due to the exclusion of several models that suggest strong wetting. The plausibility of favourable conditions in the future for environmental flows in the Rufiji also reduces after sub-setting in W2 and W4. Overall, we can detect three types of risk profile; (i) defined by low risk but with a few plausible high-risk futures, as for hydropower in both basins; (ii) with risk rather equally spread; and (iii) where risk is concentrated, as with irrigation in the Rufiji.

We note that the risk profile is clearly more sensitive to the climate model bias correction approach (DF vs QM) than to model weighting. For a given sector, the difference in risk profile is generally greater between Figs. 5 and 6 than across the weighting methods in either Figs. 5 or 6. The bias correction methods result in differing space-time structures of projected precipitation changes and accordingly hydrological responses. Broadly, the QM simulations show greater non-linearity between changes in precipitation and river flow, resulting in more dramatic risk profiles. This is likely due to the QM method retaining projected changes to weather/climate variability which are ignored in the simple DF approach ((the physical mechanisms underlying these precipitation and hydrological responses are examined in a companion paper).

4 Discussion and conclusions

Climate risk assessment underpins climate change adaptation planning and risk management activities, particularly for 'high stakes' and potentially irreversible decisions. Such assessments are usually founded on climate model projections. Often the CMIP 'ensemble of opportunity' is the starting point for this exercise, with the spread of model projections assumed to be the simplest estimate of projection uncertainty. Such uncertainty in the projected future climates remains stubbornly high for many key variables at the spatial scales of sectoral climate impacts (typically local to sub-national). This has led to the development of approaches of DMUCU. At the same time, evaluation of climate models has revealed many sources of model bias and error and recognition that models differ markedly in their performance. As a result, methods have been devised to weight models in the CMIP ensemble according to their performance and the extent to which this constrains uncertainty has been assessed but typically only for simple illustrative climate indicators (like mean changes).

In the absence of agreed good practice on model weighting, this paper examines the extent to which various model performance-weighting approaches modulate the risk profile of decision-relevant metrics, i.e. might model weighting actually matter? We use two case studies illustrative of water resource management across the water-energy-food nexus. In short, our case-specific analysis suggests that model weighting has little systematic effect on risk profiles of indicators of sectoral investment failure. Sensitivity to climate model weighting is very small in comparison to overall uncertainty and is considerably less than the uncertainty introduced by the choice of model bias correction approach, for example. Overall, the use of a more sophisticated quantile mapping bias correction to climate models tends to flatten the projected risk profiles. This increase in projection uncertainty is an unsurprising outcome of incorporating changes to climate variability as well as mean climate. Of course, there are instances in which the more aggressive model weighting approaches modifies the plausibility of higher or damaging risk and this could potentially be important to investment decisions. However, given the likely strong sensitivity of weighting to the choice of performance weights

used, the first order indication here is that such an exercise is not likely to be especially fruitful in our cases.

In this paper, we consider only risk assessment. We can conclude that given lack of clarity on good practice in model weighting, its high technical demands, considerable inputs from highly skilled scientists and the rather limited effect of weighting on impact risk profiles; there may be little incentive to invoke such approaches in real-world decision making. However, this research is not designed to directly inform real-world decisions, and any interpretation of the various risk profiles would require thorough engagement with stakeholders. Indeed, a more extensive analysis of the effects of alternative water resource performance metrics and the relative importance of the water, energy and food sectors could yield different findings. In future work, we are testing the sensitivity of robust sectoral investment portfolios defined across multiple performance-criteria to climate model weighting.

We recognize these findings are specific to the present case studies. Applying similar methods to other regions could produce different results due to strongly differing degrees and nature of (i) future climate uncertainty, (ii) model performance and (iii) baseline risk (which in our cases is quite low). We note a number of important assumptions and restrictions in our analysis, identified in the “Methods” sections, which raise caveats around too strong an interpretation or extrapolation of the findings here. The strong sensitivity of the effect of weighting on uncertainty and risk profiles to the choice of climate model performance metrics remains a major issue and one which requires further attention. We make no judgement here about the relative merits of the weighting approaches applied.

The full climate risk profile is important in advising policy makers and guiding further research; the most serious potential hydropower risk in the Rufiji basin (e.g. over 35%) occurs with just three outlier models, which, depending on stakeholder risk appetite, could be further scrutinized for their representation of regional climate drivers and teleconnections and discounted if found unsatisfactory, to avoid making expensive modifications on the basis of unreliable information. Other risk profiles are either characterised by a concentration of models around a certain risk or by an even spread from low to high risk, such as for the environmental flow indicators, with uncertainty less likely to be easily constrained.

It is clear that making the step forward from the assumption of ensemble model ‘democracy’ is justified based on our understanding of model performance (Eyring et al. 2019), and as such is probably desirable in applications. Our analysis is cautionary in that context. Nevertheless, advances in the science of process-based model evaluation, especially focussing on identifying relevant ‘emergent constraints’ on model performance raises the possibility of more robust model weighting approaches with which to enhance the credibility of model projections.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10584-021-02991-8>.

Acknowledgements This work was carried out under the Future Climate for Africa UMFULA project, with financial support from the UK Natural Environment Research Council (NERC), grant refs: NE/M020258, NE/M020398/1 see <http://www.futureclimateafrica.org/>, last access: 01 July 2020 and the UK Government’s Department for International Development (DfID).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and

indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Baumberger C, Knutti R, Hadom GH (2017) Building confidence in climate model projections: an analysis of inferences from fit. *Wiley Interdiscip Rev Clim Chang*. <https://doi.org/10.1002/wcc.454>
- Bhave AG, Conway D, Dessai S, Stainforth DA (2018) Water resource planning under future climate and socioeconomic uncertainty in the Cauvery River basin in Karnataka, India. *Water Resour Res* 54:708–728. <https://doi.org/10.1002/2017wr020970>
- Bhave AG, Bulcock L, Dessai S, Conway D, Jewitt G, Dougill AJ, Kolusu SR, Mkwambisi D (2020) Lake Malawi's threshold behaviour: a stakeholder-informed model to simulate sensitivity to climate change. *J Hydrol* 584:124671. <https://doi.org/10.1016/j.jhydrol.2020.124671>
- Conway D, Nicholls RJ, Brown S, Tebboth MGL, Adger WN, Ahmad B, Biemans H, Crick F, Lutz AF, Campos RSD, Said M, Singh C, Zaroug MAH, Ludi E, New M, Wester P (2019) The need for bottom-up assessments of climate risks and adaptation in climate-sensitive regions. *Nat Clim Chang* 9:503–511. <https://doi.org/10.1038/s41558-019-0502-0>
- Déqué M, Somot S (2010) Weighted frequency distributions express modelling uncertainties in the ENSEMBLES regional climate experiments. *Clim Res* 44:195–209. <https://doi.org/10.3354/cr00866>
- Duvail S, Hamerlynck O (2007) The Rufiji River flood: plague or blessing? *Int J Biometeorol* 52:33–42. <https://doi.org/10.1007/s00484-007-0105-8>
- Erfani T, Pachos K, Harou JJ (2018) Real-options water supply planning: multistage scenario trees for adaptive and flexible capacity expansion under probabilistic climate change uncertainty. *Water Resour Res* 54(7): 5069–5087
- Eyring V, Cox PM, Flato GM, Gleckler PJ, Abramowitz G, Caldwell P, Collins WD, Gier BK, Hall AD, Hoffman FM, Hurtt GC, Jahn A, Jones CD, Klein SA, Krasting JP, Kwiatkowski L, Lorenz R, Maloney E, Meehl GA, Pendergrass AG, Pincus R, Ruane AC, Russell JL, Sanderson BM, Santer BD, Sherwood SC, Simpson IR, Stouffer RJ, Williamson MS (2019) Taking climate model evaluation to the next level. *Nat Clim Chang* 9:102–110. <https://doi.org/10.1038/s41558-018-0355-y>
- Famien AM, Janicot S, Ochou AD, Vrac M, Defrance D, Sultan B, Noël T (2018) A bias-corrected CMIP5 dataset for Africa using the CDF-t method – a contribution to agricultural impact studies. *Earth Syst Dynam* 9:313–338. <https://doi.org/10.5194/esd-9-313-2018>
- Geressu R, Siderius C, Harou JJ, Kashaigili J, Pettinotti L, Conway D (2020) Assessing river basin development given water-energy-food-environment interdependencies. *Earth's Future* 8(8):e2019EF001464
- Gershunov A, Shulgina T, Clemesha RE, Guirguis K, Pierce DW, Dettinger MD, Lavers DA, Cayan DR, Polade SD, Kalansky J, Ralph FM (2019) Precipitation regime change in Western North America: the role of atmospheric rivers. *Sci Rep* 9(1):1–11. <https://doi.org/10.1038/s41598-019-46169-w>
- Harris I, Jones P, Osborn T, Lister D (2013) Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 dataset. *Int J Climatol* 34:623–642. <https://doi.org/10.1002/joc.3711>
- Harrison L, Funk C, McNally A, Shukla S, Husak G (2019) Pacific sea surface temperature linkages with Tanzania's multi-season drying trends. *Int J Climatol* 39:3057–3075. <https://doi.org/10.1002/joc.6003>
- Hewitson B, Waagsaether K, Wohland J, Kloppers K, Kara T (2017) Climate information websites: an evolving landscape. *Wiley Interdiscip Rev Clim Chang* 8(5):e470
- Hunt A, Watkiss P (2011) Climate change impacts and adaptation in cities: a review of the literature. *Clim Chang* 104(1):13–49
- Hurford AP, Harou JJ, Bonzanigo L, Ray PA, Karki P, Bharati L, Chinnasamy P (2020) Efficient and robust hydropower system design under uncertainty—a demonstration in Nepal. *Renewable and Sustainable Energy Reviews*, 132, p.109910. Hunt, A., Watkiss, P. (2011) Climate change impacts and adaptation in cities: a review of the literature. *Clim Chang* 104:13–49. <https://doi.org/10.1007/s10584-010-9975-6>
- Intergovernmental Panel on Climate Change (2013) Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, et al., editors Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Available: http://www.climatechange2013.org/images/uploads/WGI_AR5_SPM_brochure.pdf

- Knutti R (2008) Should we believe model predictions of future climate change? *Philos Trans R Soc A Math Phys Eng Sci* 366(1885):4647–4664
- Knutti R, Furrer R, Tebaldi C, Cernak J, Meehl GA (2010) Challenges in combining projections from multiple climate models. *J Clim* 23:2739–2758. <https://doi.org/10.1175/2009jcli3361.1>
- Kolusu SR, Shamsudduha M, Todd MC, Taylor RG, Seddon D, Kashaigili JJ, Ebrahim GY, Cuthbert MO, Sorensen JPR, Villholth KG, Macdonald AM, Macleod DA (2019) The El Niño event of 2015–2016: climate anomalies and their impact on groundwater resources in East and Southern Africa. *Hydrol Earth Syst Sci* 23:1751–1762. <https://doi.org/10.5194/hess-23-1751-2019>
- Lankford B, Beale T (2007) Equilibrium and non-equilibrium theories of sustainable water resources management: dynamic river basin and irrigation behaviour in Tanzania. *Glob Environ Chang* 17:168–180. <https://doi.org/10.1016/j.gloenvcha.2006.05.003>
- Luhunga PM, Kijazi AL, Chang'a L, Kondowe A, Ng'ongolo H, Mtongori H (2018) Climate change projections for Tanzania based on high-resolution regional climate models from the Coordinated Regional Climate Downscaling Experiment (CORDEX)-Africa. *Frontiers in Environmental Science* 6:122
- Marchau VA, Walker WE, Bloemen PJ, Popper SW (2019) Decision making under deep uncertainty: from theory to practice. Springer Nature, p. 405 <https://doi.org/10.1007/978-3-030-05252-2>
- Masson D, Knutti R (2011) Spatial-scale dependence of climate model performance in the CMIP3 ensemble. *J Clim* 24:2680–2692. <https://doi.org/10.1175/2011jcli3513.1>
- Massoud E, Espinoza V, Guan B, Waliser D (2019) Global climate model ensemble approaches for future projections of atmospheric rivers. *Earth's Future* 7:1136–1151. <https://doi.org/10.1029/2019ef001249>
- Massoud, E.C., Lee, H., Gibson, P.B., Loikith, P. and Waliser, D.E (2020) Bayesian model averaging of climate model projections constrained by precipitation observations over the contiguous United States. *J Hydrometeorol*, 21(10), pp.2401–2418
- Munday C, Washington R (2019) Controls on the diversity in climate model projections of early summer drying over Southern Africa. *J Clim* 32:3707–3725. <https://doi.org/10.1175/jcli-d-18-0463.1>
- Nissan H, Goddard L, Perez ECD, Furlow J, Baethgen W, Thomson MC, Mason SJ (2019) On the use and misuse of climate change projections in international development. *Wiley Interdiscip Rev Clim Chang*. <https://doi.org/10.1002/wcc.579>
- Pennell C, Reichler T (2011) On the effective number of climate models. *J Clim* 24:2358–2367. <https://doi.org/10.1175/2010jcli3814.1>
- Ray PA, Brown CM (2015) Confronting climate uncertainty in water resources planning and project design: the decision tree framework. World Bank, Washington, DC
- Rowell DP (2019) An observational constraint on CMIP5 projections of the east African long rains and southern Indian Ocean warming. *Geophys Res Lett* 46:6050–6058. <https://doi.org/10.1029/2019gl082847>
- Rowell DP, Chadwick R (2018) Causes of the uncertainty in projections of tropical terrestrial rainfall change: East Africa. *J Clim* 31(15):5977–5995. <https://doi.org/10.1175/JCLI-D-17-0830.1>
- Rowell DP, Senior CA, Vellinga M, Graham RJ (2016) Can climate projection uncertainty be constrained over Africa using metrics of contemporary performance? *Clim Chang* 134:621–633. <https://doi.org/10.1007/s10584-015-1554-4>
- Sanderson BM, Wehner M, Knutti R (2017) Skill and independence weighting for multi-model assessments. *Geosci Model Dev* 10:2379–2395. <https://doi.org/10.5194/gmd-10-2379-2017>
- Schneider U, Finger P, Meyer-Christoffer A, Rustemeier E, Ziese M, Becker A (2017) Evaluating the hydrological cycle over land using the newly-corrected precipitation climatology from the Global Precipitation Climatology Centre (GPCC). *Atmosphere* 8:52. <https://doi.org/10.3390/atmos8030052>
- Seaby L, Refsgaard J, Sonnenborg T, Stisen S, Christensen J, Jensen K (2013) Assessment of robustness and significance of climate change signals for an ensemble of distribution-based scaled climate projections. *J Hydrol* 486:479–493. <https://doi.org/10.1016/j.jhydrol.2013.02.015>
- Siderius C, Biemans H, Kashaigili JJ, Conway D (2018) Going local: evaluating and regionalizing a global hydrological model's simulation of river flows in a medium-sized East African basin. *Journal of Hydrology: Regional Studies* 19:349–364. <https://doi.org/10.1016/j.ejrh.2018.10.007>
- Stainforth D, Allen M, Tredger E, Smith L (2007) Confidence, uncertainty and decision-support relevance in climate predictions. *Philos Trans R Soc A Math Phys Eng Sci* 365:2145–2161. <https://doi.org/10.1098/rsta.2007.2074>
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *Bull Amer Meteor Soc* 93:485–498. <https://doi.org/10.1175/BAMS-D-11-00094.1>
- Teutschbein C, Seibert J (2012) Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. *J Hydrol* 456–457:12–29. <https://doi.org/10.1016/j.jhydrol.2012.05.052>

- Weaver CP, Lempert RJ, Brown C, Hall JA, Revell D, Sarewitz D (2013) Improving the contribution of climate model information to decision making: the value and demands of robust decision frameworks. *Wiley Interdiscip Rev Clim Chang* 4(1):39–60
- WREM-International (2015) WREM-International Rufiji IWRMD Plan Draft Final Report. Volume I, Report Prepared for the United Republic of Tanzania, ministry of Water WREM International Inc., Atlanta, Georgia, USA

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.